

# POL-index

## Polska Baza Cytowań

Wojciech Fenrich  
[w.fenrich@icm.edu.pl](mailto:w.fenrich@icm.edu.pl)



# Geneza

- **"Brak PIF"** - czysto formalna ocena czasopism humanistycznych i społecznych w części B wykazu czasopism punktowanych.
- **Cel:** stworzyć narzędzie pozwalające uzyskać sensowne informacje o wpływie polskich czasopism humanistycznych i społecznych.

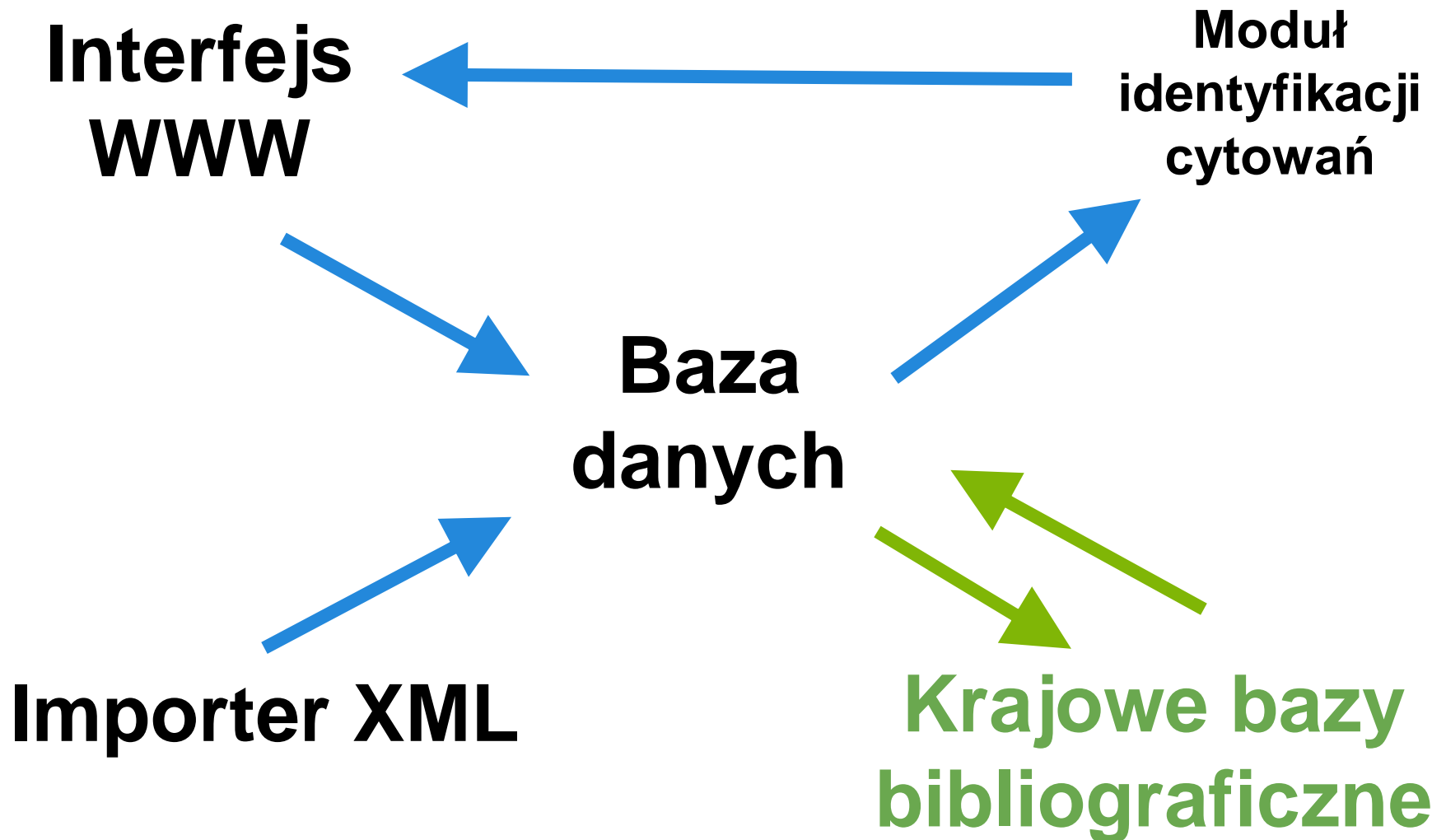


# Najważniejsze problemy

- **Skąd dane?** - Konieczny elastyczny model pozyskiwania danych na potrzeby POL-index, minimalny zakres danych.
- **Czy elastyczność nie oznacza niskiej jakości danych?** – Baza otwarta, możliwość zgłaszania uwag.
- **Krótki termin realizacji** – Nie wszystkie funkcje pojawią się od razu.



# Elementy systemu POL-index



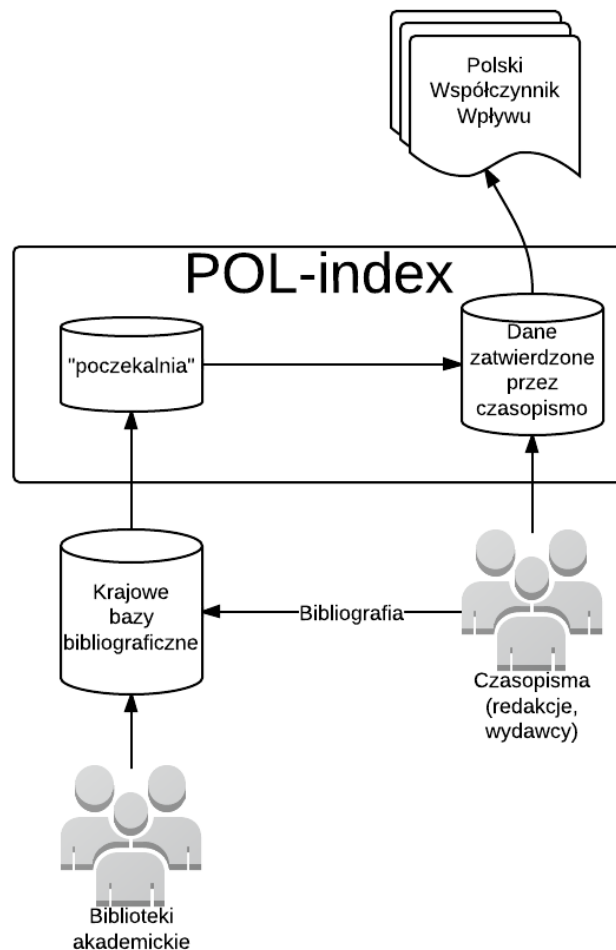
# Przepływ danych

## – wersja minimalna

- Dane są wprowadzane i zatwierdzane przez przedstawicieli czasopism w ramach procesu ewaluacji czasopism (nie jest istotny styl cytowania, ew. konieczność opracowania przypisów dolnych) – obowiązek lub zachęta.
- Wstępna maszynowa identyfikacja cytowań.
- Etap zgłaszania uwag i ich weryfikacji (rzetelność danych, poprawność identyfikacji cytowań).
- Ostateczna maszynowa identyfikacja cytowań.
- Prezentacja wyników, obliczenie PWW.



# Przepływ danych – wersja pośrednia



# Przepływ danych – wersja optymalna

- Wszystkie czasopisma z listy B indeksowane są w bazach współpracujących z POL-index - czasopisma nie muszą uzupełniać i weryfikować danych.
- Współpraca z wydawnictwami naukowymi - indeksacja książek.
- Rozwój algorytmów identyfikacji cytowań – minimalizacja potrzeby weryfikacji danych.



# Jak liczyć Polski Współczynnik Wpływu (PWW)? – Najważniejsze problemy

1. PWW jako element oceny czasopism powinien być zrozumiały.
2. Inna struktura cytowań w czasopismach humanistycznych i społecznych.
3. Dłuższy czas od momentu publikacji artykułu do momentu jego zacytowania w przypadku nauk społecznych i humanistycznych. **Ale...**
4. PWW powinien odzwierciedlać **aktualny** stan czasopisma – nie można kierować analiz zbyt daleko w przeszłość.





# Przepływ danych i problemy warunkują zakres danych

- **Jakie czasopisma?** – Przede wszystkim czasopisma składające ankietę (listy B), również zainteresowane czasopisma z list A i C.
- **Zakres czasowy:** w pierwszym roku – pełne rekordy za dwa poprzednie lata, rekordy skrócone od trzeciego do szóstego roku wstecz. W kolejnych latach – tylko pełne rekordy z roku poprzedniego.
- **Minimalne dane o artykule:** nazwa czasopisma\*, tom\*, zeszyt\*, tytuł artykułu, strony, imiona i nazwiska autorów, afiliacje autorów (nieobowiązkowe), bibliografia załącznikowa (tylko pełne rekordy).

# Propozycja sposobu wyliczania PWW – jak uporządkować czasopisma?

$$\frac{C_{(S-5)-(S-1)}^S}{N_{(S-5)-(S-1)}} + \frac{C_{(S-6)-(S-2)}^{S-1}}{N_{(S-6)-(S-2)}}$$

- **Pierwszy składnik sumy:** Liczba cytowań w roku S do wszystkich artykułów opublikowanych w czasopiśmie w latach od S-5 do S-1, podzielona przez liczbę cytowalnych artykułów naukowych opublikowanych w czasopiśmie w latach od S-5 do S-1.
- **Drugi składnik sumy:** jak wyżej, dla roku S-1.
- Pełne rekordy – 2 lata. Rekordy skrócone – kolejne 5 lat.



# PWW - dlaczego tak?

- Czasopisma są oceniane za ostatnie 2 lata – podstawowa zasada ewaluacji.
- PWW stara się pogodzić dłuższy czas oczekiwania na cytowania z koniecznością odzwierciedlenia aktualnego stanu czasopisma.
- Większa stabilność punktacji - co roku zmienia się jedynie część danych.

# POL-index dziś

- Baza danych, importer XML, interfejs WWW do wprowadzania danych, moduł analizy cytowań - końcowe etapy prac.
- Kiedy start? – Ostateczna decyzja zależy od decyzji Ministerstwa Nauki.



# Przyszłość?

- Rozwój interfejsu WWW: część analityczna.
- Dalsze doskonalenie algorytmów automatycznego rozpoznawania cytowań.
- W stronę pośredniego, a następnie optymalnego przepływu danych.
- Indeksacja monografii naukowych.



# Dziękuję za uwagę

Wojciech Fenrich  
w.fenrich@icm.edu.pl

